<Bioinformatics Center> Biological Information Networks

Bioinformatics Center
- Biological Information Networks -

http://www.bic.kyoto-u.ac.jp/takutsu/index.html

Prof
AKUTSU, Tatsuya
(D Eng)

Assist Prof
UEDA, Nobuhisa
(D Eng)

Assist Prof
HAYASHIDA, Morihiro
(D Inf)

Guest Res Assoc
NACHER, Jose C
(Ph D)

PD
TAMURA, Takeyuki
(D Inf)

Students

MOURI, Kazunari (D1)
TAKEMOTO, Kazuhiro (D1)
KASHIMA, Hisashi (D1)
BROWN, John (M2)
OOTAKA, Ryu (M2)
SAKAI, Daisuke (M2)
TAKEUCHI, Shigeki (M2)
POOLSAP, Unyanee (M1)
CHO, Uncho (M1)
URATA, Takashi (M1)

Visitors

Ms ZHANG, Shuqin — The University of Hong Kong, China, 1 April–30 June 2006
Prof SMOLA, Alex J — National Information and Communications Technology, Australia, 13 October 2006

Scope of Research

Due to rapid progress of the genome projects, whole genome sequences of organisms ranging from bacteria to human have become available. In order to understand the meaning behind the genetic code, we have been developing algorithms and software tools for analyzing biological data based on advanced information technologies such as theory of algorithms, artificial intelligence, and machine learning. We are recently studying the following topics: systems biology, scale-free networks, protein structure prediction, inference of biological networks, chemo-informatics, discrete and stochastic methods for bioinformatics.

Research Activities (Year 2006)

Presentations

On the Complexity of Finding Control Strategies for Boolean Networks, Akutsu T, Hayashida M, Ching W-K, Ng M, The 4th Asia-Pacific Bioinformatics Conference, 15 February.

Topological Aspects of Protein Networks, Nacher JC, Hayashida M, Akutsu T, Workshop on Emergent Intelligence on Networked Agents, 8 May.

Theoretical and Computational Analyses of Structures of Metabolic Networks and Protein-protein Interaction Networks, Akutsu T, Nacher JC, The First International Conference on Computational Systems Biology, 21 July.

Multiple Methods for Protein Side Chain Packing Using Maximum Weight Cliques, Brown JB, K.C. D, Tomita E, Akutsu T, The 6th International Workshop on Bioinformatics and Systems Biology, 24 July.

Identification of Metabolic Units Induced by Environmental Signals, Nacher JC, Schwartz J-M, Kanehisa M, Akutsu T, The 14th Annual International Conference on Intelligent Systems for Molecular Biology, 7 August.

Grants

Akutsu T, Miyano S, Maruyama O, Ueda N, Algorithms for Extracting Common Patterns from Structured Biological Data, Grant-in-Aid for Scientific Research (B), 1 April 2004–31 March 2008.

Akutsu T, Goto S, Mochizuki A, Tokita K, Mathematical Analysis of Structure and Dynamics of Biological Information Networks, Grant-in-Aid for Priority Area Research, 1 April 2005–31 March 2010.

Ueda N, Statistical Language Models That Generate a Pair of Sequences for Sequence Analysis, Grant-in-Aid for Encouragement of Young Scientists, 1 April 2003–31 March 2006.

Identification of Metabolic Units Induced by Environmental Signal

Biological cells continually need to adapt the activity levels of metabolic functions to changes in their living environment. Although genome-wide transcriptional data have been gathered in a large variety of environmental conditions, the connections between the expression response to external changes and the induction or repression of specific metabolic functions have not been investigated at the genome scale.

We present here a correlation-based analysis for identifying the expression response of genes involved in metabolism to specific external signals, and apply it to analyze the transcriptional response of Saccharomyces cerevisiae to different stress conditions. We show that this approach leads to new insights about the specificity of the genomic response to given environmental changes, and allows us to identify genes that are particularly sensitive to a unique condition. We then integrate these signal-induced expression data with structural data of the yeast metabolic network and analyze the topological properties of the induced or repressed sub-networks. They reveal significant discrepancies from random networks, and in particular exhibit a high connectivity, allowing them to be mapped back to complete metabolic routes.

Nacher J.C., Schwartz J-M., Kanehisa M., Akutsu T., *Bioinformatics*, **22**, e375-e383 (2006).

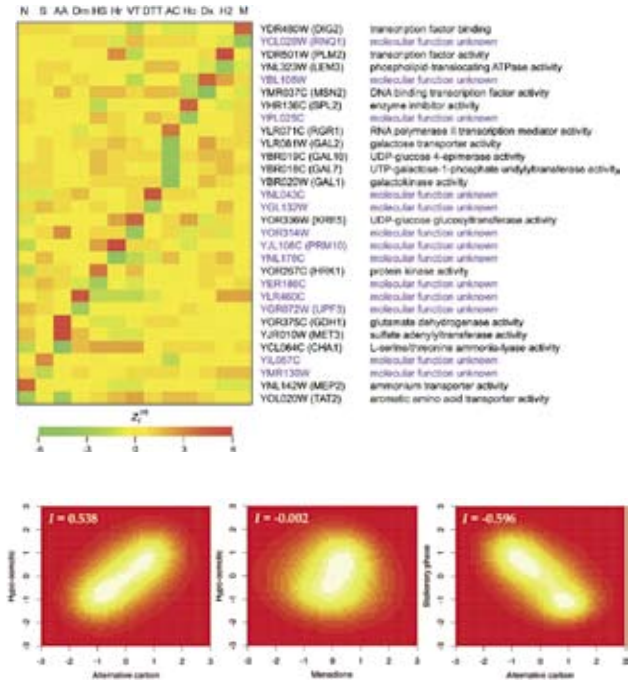

Figure 1. (Top) Genes with a unique response to specific signals. (Bottom) Representative examples of density distributions of pairs of z_i (s) scores.

A Novel Clustering Method for Analysis of Biological Networks Using Maximal Components of Graphs

Clustering is one of fundamental techniques in bioinformatics. Indeed, many clustering methods have been developed and/or applied for analyzing various kinds of biological data. However, these clustering methods such as widely used linkage methods are based on similarities between two elements or two clusters, and relations with other elements or clusters are not so much taken into account. Relations between biological entities are often represented as networks or (almost equivalently) graphs. Since such networks are considered to have much information, clustering based on network structures might be useful. The set of all maximal components of a graph essentially contains all information on minimum cuts for all pairs of nodes, where a maximal component is a set of nodes with high connectivity. It is known that a set of maximal components constitutes a laminar structure, which is essentially a hierarchical structure. Thus, we develop a novel clustering method using maximal components for an undirected network. In this method, nodes are partitioned into clusters by selecting disjoint maximal components. In this study, we apply the proposed method to clustering of protein sequences and compare with the single-linkage, complete-linkage and average-linkage methods. We evaluate the computed clusters using *P*-values for GO (GeneOntology) terms. The results suggest the effectiveness of the proposed method.

Hayashida M., Akutsu T., Nagamochi H., *Proc. 5th Asia-Pacific Bioinformatics Conference*, in press.

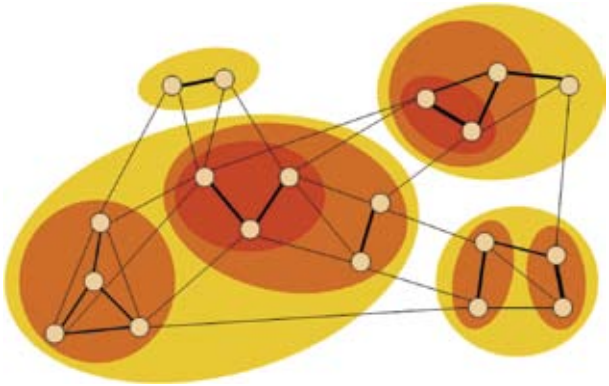

Figure 2. An example of maximal components of a graph.